



King's Research Portal

Document Version
Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Steves, C. J., Menni, C., Spector, T. D., Small, K. S., Christofidou, P., Williams, F. M. K., Hammond, C. J., Hart, D. J., & Bowyer, R. C. E. (Accepted/In press). TwinsUK: The UK Adult Twin Registry Update . *TWIN RESEARCH AND HUMAN GENETICS*.

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Title: TwinsUK: The UK Adult Twin Registry Update

Running title: The TwinsUK Registry Update

Author List: Serena Verdi¹, Golboo Abbasian¹, Ruth C. E Bowyer¹, Genevieve Lachance¹, Darioush Yarand¹, Paraskevi Christofidou¹, Massimo Mangino¹, Cristina Menni¹, Jordana T. Bell¹, Mario Falchi¹, Kerrin S. Small¹, Frances M. K Williams¹, Christopher J. Hammond¹, Deborah J. Hart¹, Timothy D. Spector¹, Claire J. Steves^{1,2}

¹ The Department of Twin Research, King's College London, St Thomas' Hospital, London, SE1 7EH, United Kingdom.

² Clinical Age Research Unit, Department of Clinical Gerontology, King's College Hospital, NHS Foundation Trust, London, United Kingdom.

Corresponding author: Dr Claire Steves: Claire.j.steves@kcl.ac.uk +44 (0) 207 188 5396

Keywords: TwinsUK, ageing, omics, microbiome, pain, ophthalmology, epidemiology

ORCID's:

Serena Verdi - <https://orcid.org/0000-0002-2421-4818>

Golboo Abbasian – <https://orcid.org/0000-0001-9767-7509>

Ruth C.E Bowyer- <https://orcid.org/0000-0002-6941-8160>

Genevieve Lachance - <https://orcid.org/0000-0003-2507-5802>

Darioush Yarand - <https://orcid.org/0000-0003-3570-3921>

Paraskevi Christofidou- <https://orcid.org/0000-0003-1223-9636>

Massimo Mangino- <https://orcid.org/0000-0002-2167-7470>

Cristina Menni - <https://orcid.org/0000-0001-9790-0571>

Jordana T. Bell- <https://orcid.org/0000-0002-3858-5986>

Mario Falchi- <https://orcid.org/0000-0002-5646-1004>

Kerrin S. Small- <https://orcid.org/0000-0003-4566-0005>

Frances M. K Williams- <https://orcid.org/0000-0002-2998-2744>

Christopher J. Hammond- <https://orcid.org/0000-0002-3227-2620>

Deborah J. Hart- <https://orcid.org/0000-0002-9066-4614>

Abstract

TwinsUK is the largest cohort of community dwelling adult twins in the UK. The registry comprises over 14,000 volunteer twins (14,838 including mixed, single and triplets), it is predominantly female (82%) and middle aged (mean age 59). In addition, over 1,800 parents and siblings of twins are registered volunteers. During the last 25 years, TwinsUK has collected numerous questionnaire responses, physical/cognitive measures and biological measures on over 8,500 subjects. Data were collected alongside four comprehensive phenotyping clinical visits to the Department of Twin Research and Genetic Epidemiology, King's College London. Such collection methods have resulted in very detailed longitudinal clinical, biochemical, behavioural, dietary and socio-economic cohort characterisation; it provides a multidisciplinary platform for the study of complex disease during the adult life-course, including the process of healthy ageing. The major strength of TwinsUK is the availability of several 'omic' technologies for a range of sample types from participants, which includes genome-wide scans of single nucleotide variants, next-generation sequencing, metabolomic profiles, microbiomics, exome sequencing, epigenetic markers, gene expression arrays, RNA sequencing and telomere length measures. TwinsUK facilitates and actively encourages sharing the 'TwinsUK' resource with the scientific community – interested researchers may request data via the TwinsUK website (<http://twinsuk.ac.uk/resources-for-researchers/access-our-data/>) for their own use or future collaboration with the study team. In addition, further cohort data collection is planned via the Wellcome Open Research gateway (<https://wellcomeopenresearch.org/gateways>). The current article presents an up-to-date report on the application of technological advances, new study procedures in the cohort and future direction of TwinsUK.

Introduction

TwinsUK is the largest adult twin registry in the UK, and is one of the most deeply phenotyped and genotyped datasets in the world. It provides a multidisciplinary platform to research both health and social related questions; with the overarching aim of understanding the aetiology of complex disease

and the ageing process. The registry was started in 1992, with the initial intention to investigate osteoporosis and osteoarthritis. Such conditions are highly prevalent in women and consequently several hundred middle-aged women were recruited and formed the core of the initial register. Success from these early studies led to a rapid expansion of TwinsUK and to date the cohort consists of 14,000 community dwelling twins, male and female, aged over 18. Current research areas of interest include the genetics of metabolic syndrome, cardiovascular disease, the musculoskeletal system, sensory impairment and ageing, as well as how the microbiome affects human health. Details of the registry's progression have been described previously (Moayyeri et al., 2013; Spector & Williams, 2006). To date, the TwinsUK registry has contributed to over 850 publications and 800 international collaborations. More detailed description of research outputs may be accessed through the study website: <http://www.twinsuk.ac.uk>

The Collection

Over the last 27 years the TwinsUK registry has been enhanced through over 80 studies, some which have been repeated over time. This has resulted in clinically rich, longitudinal phenotype information, (Table 2), which may be categorised to 4 distinct time points (Verdi et al., 2019). Recruitment strategies have predominantly involved media campaigns. These have offered opportunities for adult twin pairs to join the registry and participate in unspecific research investigating various common diseases, without selecting for particular diseases or traits. At baseline (1992-2004), over 7,000 twins responded to annual questionnaires and approximately 5,500 twins attended a full comprehensive clinical visit, which included several project-led studies. Age-matched characteristics of these volunteer twins were found not to differ from a singleton population-based cohort of British women (Chingford study) (Andrew et al., 2001), apart from a life-long lower weight in MZ twins of approximately 1kg. The follow-up visit occurred between April 2004 and May 2007, in which 3,725 twins in the registry attended a 1-day clinical visit and an additional 1,299 twins posted blood taken via their GPs for DNA sampling. Participants ranged aged between 18 and 82 years (mean 52.5 ± 13

years) and the majority were female (89%). Protocols for the baseline and initial follow-up visit have been described previously (Spector & Williams, 2006).

The second wave of follow-up visits (August 2007- April 2012) aimed to investigate the ageing process; HATS (Healthy Ageing Twin Study). Inclusion criteria were women aged ≥ 40 years with at least one previous clinical visit ($n=4,610$). In total, 3,125 women (mean age 59.6 ± 9 years) attended the clinical visit. Follow-up time between first and last visits ranged between 6.1 - 17.4 years, with over 600 of the participants having 4 or more previous clinical visits. HATS outcomes have previously been described (Moayyeri et al., 2012) including details of data collection (Moayyeri et al., 2013).

The 3rd wave of follow-up visits (May 2012 and May 2018) was performed to understand the interactions in disease processes between genes and the environment, as part of the Biomedical Research Centre (BRC) study. All participants of the TwinsUK registry were invited to attend a comprehensive clinical visit, which included collection of bone density/whole body scan, cognitive and lung function, hearing and eye tests, fitness assessment (gait speed, chair stands and grip strength) and collection of blood, urine, stool and salivary samples. In total 6,686 clinical visits were made, with 3,620 volunteers attending at least once and 1,531 volunteers attended the clinic on two occasions with an average of 4 years between visits. In addition to the clinical visit, 6,300 questionnaires were returned, complementing clinical data collected during the visit. Since February 2019, a further wave of follow up visits has commenced which aims to continue the longitudinal data collection, and adds further dynamic phenotyping and blood measurement over a six-hour visit incorporating standardised meals.

Table 1. TwinsUK Registry Update

Register Name	TwinsUK
Country	United Kingdom

Ascertainment type	Unselected Volunteers
Opposite sex twins	Yes
Number of pairs n separated by year of birth and sex)	1900- 1909: 2 FF 1910- 1919: 14 FF 1920-1929: 304 FF; 48 MM; 4 FM 1930- 1939: 1,418 FF; 196 MM; 36 FM 1940- 1949: 2,352 FF; 312 MM; 64 FM 1950- 1959: 2,380 FF; 486 MM; 98 FM 1960- 1969: 2,104 FF; 502 MM; 70 FM 1970-1979: 1,702 FF; 506 MM; 56 FM 1980-1989: 950 FF; 186 MM; 58 FM 1990- 1999: 554 FF; 162 MM; 56 FM 2000- 2001: 50 FF; 12 MM; 4 FM
Total	11, 830 FF; 2,410 MM, 446FM 14,686 (7,343 pairs)
Zygosity Classification	51% monozygotic (MZ); 49% dizygotic (DZ). Assessed by the ‘peas in a pod’ questionnaire during registration (Jarrar et al. 2018) and confirmed via genotyping / genome-wide association studies.
Major interests	Common complex diseases, ageing traits, nutrition and microbiome.
Traits measured	Questionnaires and clinical examinations for an extensive range of traits, within the categories: twinship & family, demographics, behaviour: dietary, sleep, physical activity, smoking, alcohol use; healthcare. Health: allergy and respiratory conditions; rheumatology; cardiology; operations, gastroenterology; oncology; urology and nephrology; neurology, psychiatry and mental health; dermatology; ophthalmology; oral health, hearing; endocrinology; obstetrics, gynaecology and fertility, wellbeing; medication histories.

Blood Samples	8332 DNA; 8726 serum; 8614 plasma; 5685 peripheral blood lymphocytes.
Other samples	8807 urine; 3206 saliva; 4576 stool; 1095 fat, muscle and skin biopsies; 5204 hair; 2711 peripheral blood mononuclear cell's; 282 TRUculture (whole blood immune cell culture system).
Comments on omics	5,654 genome wide association data 2,000 (UK10K), 2,377 (HLI) next gen sequenced data 5,000 (Epitwin) DNA methylation data Data available for transcriptome across 800 individuals in 3 tissues.
Main sources of funding	The Wellcome Trust, the Medical Research Council and the Chronic Disease Research Foundation (CDRF) currently support TwinsUK. Since 2012, the GSTT/KCL NIHR Biomedical Research Centre has also provided support for twin visits. Several smaller charitable bodies and commercial companies have also contributed smaller grants.
Contact	Prof Timothy D. Spector
Address	The Department of Twin Research and Epidemiology, St. Thomas Hospital, London, SE1 7EH.
Email	tim.spector@kcl.ac.uk or Victoria.vazquez@kcl.ac.uk
Website	http://twinsuk.ac.uk/

Note: FF = female–female; MM = male–male; FM = female–male

Longitudinal data

Detailed clinical and biochemical phenotypes have been collected using harmonised protocols at each visit stage. A summary of a selection of clinical phenotypes are outlined in Table 2. In addition, questionnaire data have been collected on an annual basis and during visits, some which measure incident clinical endpoints such as cardiovascular accident, type 2 diabetes, chronic obstructive pulmonary disease, which have previously been described (Verdi et al., 2019). Three main comprehensive questionnaires (TwinsUK Baseline Health', 'Baseline Core' and 'Longitudinal Core')

were collected between 2004 and 2018 (detailed in Table 3). These were in paper format, completed at respondents address and returned to the research facility. Over 2500 participants completed all three main questionnaires and 2,300 completed either two of the main questionnaires. Furthermore, the demographic of the cohort provides an excellent resource to study ageing where longitudinal changes are important to consider. Table 4 provides summaries of the key cognitive and frailty phenotypes we have acquired to explore questions in this area.

Alongside regular visits and questionnaires, TwinsUK has data linkage to official cancer and mortality data for retrospective analysis and future follow up. Additional links to national health, education and environmental records to our own database are being established at present.

Table 2. Longitudinal Data Available in the TwinsUK Registry Participants

Phenotype	N of measurements	N of participants	N of participants with ≥ 2 measurements	N of participants with ≥ 4 measurements	Maximum number of visits	Maximum duration of follow-up (year)
Insulin	19,048	8,298	4,829	1,767	12	24.2
Glucose	23,079	9,227	5,444	2,421	13	24.7
Blood Pressure	22,252	9,422	5,405	2,299	8	25
Lipid profile	18,933	8,496	4,754	1,703	9	24.7
Respiratory function	18,665	8,880	4,790	1,494	7	24.2
Hip BMD	18,671	8,713	4,798	1,563	7	21.3
Spine BMD	18,263	8,497	4,665	1,553	7	19.5
Whole body DXA	17,690	8,148	4,566	1,529	7	21.5
Grip Strength	13,551	6,780	4,008	577	4	13.4
BMI	20, 881	8076	5288	2328	8	24.4

Note: DXA= dual-energy X-ray absorptiometry; BMD= bone mineral density; BMI = body mass index

Table 3. Longitudinal measures of main TwinsUK Questionnaires

Questionnaire	Year	N of responses	N of participants	Maximum duration of follow-up (year)
TwinsUK Baseline Health Questionnaire	2004-2014	7,000	5,400	10
Baseline Core Questionnaire	2004-2018	6,300	6,300	NA
Longitudinal Core Questionnaire	2017-2018	3,300	3,300	NA

Note: The Baseline Core and Longitudinal Core were collected 4 years apart. NA= not applicable.

Table 4. Longitudinal cognitive and frailty measures Available in the TwinsUK Registry

Participants

Phenotype	N of measurements	N of participants	N of participants with two measurements	Mean Follow-up (year)	Maximum duration of follow-up (year)
Mini-Mental State Examination	5,426	4,951	475	3.9	4.4
Standing Balance	3,941	3,938	NA	NA	2.3
Gait Speed	5,387	4,918	469	3.9	4.4
Chair stand	3,928	3,927	NA	NA	2.3
Verbal Fluency	5,413	4,941	472	3.9	4.4
Deary Liewald Reaction Time Test	3,227	3,226	NA	NA	NA

CANTAB Paired Associates Learning	2,091	2,091	NA	NA	NA
-----------------------------------	-------	-------	----	----	----

Note: CANTAB= Cambridge Neuropsychological Test Automated Battery; NA= not applicable.

Novel molecular and genetic phenotypes

In addition to epidemiological and clinical phenotypes collected from clinical visits, numerous biological samples including body fluids (blood, urine, saliva, stool, sebum) and tissue (hair follicle, colonic mucosa, fat and skin biopsy) have been generously donated. Details of the samples collected are summarised in Table 1 and (Verdi et al., 2019). Collection methods have been described in their respective research publications which can be found on the TwinsUK website (<http://www.twinsuk.ac.uk>). Here we describe the omic techniques (genome-wide association studies, epigenomics, next generation sequencing, metagenomics, metabolomics and microbiomics) which have been employed on biological samples and phenotypes from TwinsUK data. Details of some phenotypes collected prior to 2012 (e.g. telomere length) have previously been described (Moayyeri et al., 2013).

- Genome-Wide Association Studies

TwinsUK has contributed to many international consortia for genome-wide association analysis of various phenotypes (Mills & Rahal, 2018). Genome-wide scan data using two chips (Illumina HumanHap300 BeadChip and Illumina HumanHap610 QuadChip) is available for 5,654 (both monozygotic and dizygotic) twins. The data has been fully imputed using ‘1000 Genomes’ and ‘Haplotype Reference Consortium – (HRC)’ reference panels. TwinsUK is a member of many ongoing international consortia for meta-analysis of various traits such as height, BMI, lipids, obesity, blood pressure and back pain phenotypes. Some of the main publications from these collaborations can be found in the TwinsUK website. Our Genome-wide data is also being used to compile Polygenic Risk Scores to isolate loci for various traits (Mills, Barban, & Tropf, 2018).

- Epigenetic markers

The first large-scale genome-wide epigenetic assessment in TwinsUK was performed on DNA methylation patterns profiled on the Illumina HumanMethylation27 BeadChip in a whole blood sample of 172 female twins. This array examines 27,578 promoter CpG-sites that map uniquely across the genome and some of these sites were found to be associated with age and age-related phenotypes (Bell et al., 2012). Subsequently, the Illumina Infinium HumanMethylation450 BeadChip was applied to up to blood samples from up to 1,000 additional MZ and DZ twins to generate higher-resolution genome-wide DNA methylation profiles (Kurushima et al., 2019; Zhang et al., 2015), as well as in 322 skin (Roos et al., 2017) and 648 adipose (Grundberg et al., 2013) tissue biopsy samples from twins. More recently, the Illumina Infinium MethylationEPIC array is being profiled in additional blood samples from over 400 twins. Further epigenetic datasets in TwinsUK cohort have also been generated as part of the EpiTwin study (<http://www.epitwin.eu>), which in collaboration with the Beijing Genomics Institute (BGI), assayed epigenomic sequencing profiles in up 5,000 samples from twins aged 16–85 years. The results include MeDIP (Methylated DNA immunoprecipitation) sequencing profiles in whole blood samples from twins discordant and concordant for a wide variety of diseases and environmental exposures (Bell et al., 2014) (Yuan et al., 2014)(Davies et al., 2014) (Bell et al., 2016).

- Gene expression markers

During the HATS visit, 856 twins with detailed clinical profiles underwent biopsies of multiple tissues as part of the MuTHER (Multiple Tissue Human Expression Resource) project. This was a Wellcome Trust funded study designed to investigate gene expression across multiple tissues simultaneously with the aim of examining mechanisms involved in common trait susceptibility. Gene expression in three tissues and derived cells, fat, skin and lymphoblastoid cell lines (LCL) was determined using Illumina whole genome expression array (HumanHT-12 version 3) comprising 48,803 probes in three technical replicates (Grundberg et al., 2012). The same skin, fat and LCL RNA samples plus an additional 400 whole blood samples were RNA sequenced as part of the EuroBATS

project (Biomarkers of Ageing using whole Transcriptome Sequencing) a European (EU-FP7) study (Buil et al., 2015).

- Whole genome sequencing

Whole genome sequencing (WGS) of 2000 healthy, deeply phenotyped twins formed part of the UK10K project, which used state-of-the-art next-generation sequencing methods to uncover rare genetic variants associated with health and disease. The data have been used extensively to describe population structure and functional annotation of rare and low-frequency variants (The UK10K Consortium et al., 2015), further details can be accessed at: www.uk10k.org. In addition, approximately 1000 exome sequences at 30-60× depth have been ascertained as part of projects with GoT2D consortium and Pfizer Inc. More recently, WGS of >30X coverage was carried out through collaboration with Human Longevity, Inc (HLI) for 2,377 individuals from the TwinsUK cohort. DNA samples were sequenced on an Illumina HiSeqX sequencer using a 150-base paired-end single-index-read format. The data have been used to disentangle the contribution of rare variants to the blood metabolome (Long et al., 2017), and are now under investigation to identify rare variants associated with complex diseases and traits, and for the inference of structural variants.

- Metabolomics profile

Fasting circulating metabolites levels (serum/plasma) have been assessed on TwinsUK study participants using different platforms: (i) Biocrates AbsoluteIDQ (163 metabolites for 1052 twins) (Menni et al., 2013a); (ii) Metabolon Inc (Metabolon Inc, Research Triangle Park, USA) (280 known metabolites and 175 unknown metabolites for 6055 twins, 591 known and 165 unknown for 2069 twins at three time points), (iii) Nightingale Health Ltd. (27 metabolic traits, 143 metabolite concentrations, 80 lipid ratios, 3 lipoprotein particle sizes and a semi-quantitative measure of albumin for 2000 twins at three time points). Genome-wide association studies have identified several genes involved in metabolic individuality (Long et al., 2017; Shin et al., 2014), as well as many health traits (see TwinsUK publications list: <http://twinsuk.ac.uk/our-research/publications/>).

Nightingale Health Ltd. (Helsinki, Finland; previously known as Brainshake Ltd) is a targeted NMR spectroscopy platform that has been extensively applied by us and others for biomarker profiling in epidemiological studies (Barrios et al., 2018). More recently, metabolomics profiling (Metabolon, Inc) has been conducted on faecal (n=1016) (Zierer et al., 2018) and salivary samples (Nag et al., 2019).

○ **Glycans**

Glycosylation is the most common form of post-transcriptional protein modification and it is a putative mechanism in the modulation of the inflammatory response. The technology to assess glycosylation has recently become high throughput, and glycosylation of immunoglobulin G (IgG) has been measured on 4900 twins while N-glycans on human serum glycoproteins on 1800. Using this, we have found that glycans are highly heritable (Menni et al, 2013b) and we have been the first to observe a number of associations between glycans and important age-related traits (Barrios et al., 2015) (Menni et al., 2018a).

- **Microbiome**

Alongside the BRC study (3rd follow-up), over 5000 faecal samples have been collected for microbiome analysis. Twin volunteers provided stool samples, stored on-site (St Thomas' Hospital, London) at -80 °C. DNA extraction and 16S rRNA sequencing using the V4 variable region of nearly 3000 samples has been completed in collaboration with Cornell University using a multiplexed approach on the Illumina MiSeq platform. Smaller subsets of twins have also been sequenced with complementary methods by the BRC Genomics Facility at King's College London. In addition, plain saliva (700) and mid-stream urine (1600) specimens have undergone similar 16S amplicon sequencing using the same primers in collaboration with University of California San Diego and Stanford University.

Diversity metrics, taxonomic levels from genus through to phylum and relative abundances of operational taxonomic units (OTUs) have been used to assess microbiota associations within the TwinsUK data. Associations have been observed with a number of health deficits and medication

usage (Jackson et al., 2018) (Jackson et al., 2016) (Le Roy et al., 2017), age related traits including frailty and cognition (Jackson et al., 2016) (Verdi et al., 2018), amongst others (Menni et al., 2018b). In addition, microbiota associations with diet (Ni Lochlainn, Bowyer, & Steves, 2018) (Menni et al., 2017a) (Menni et al., 2017b) and socioeconomic status (Bowyer et al., 2019) have been found. More recently Amplicon Sequence Variants (ASV), also known as exact sequence, have been generated from 3045 stool samples. This approach offers a higher resolution than the OTU, allowing for greater sensitivity and specificity in identifying the taxonomic associations with traits (Wells et al., 2019).

- Metagenomics

Whole metagenomic shotgun sequencing (WMGS) has been performed on faecal samples in two batches comprising 250 and 1,004 volunteers from the TwinsUK registry. This larger dataset, including 161 monozygotic twin pairs, 201 dizygotic twin pairs, and 280 singletons generated an average of 39M high-quality microbial reads per sample. Taxonomic and functional information have been inferred from the WGMS data. These results are being studied to determine the influence of the microbiome on the faecal and host metabolome, and to identify bacterial species and function mediating microbiome-associated increased risk for common disease.

Dietary Phenotypes

TwinsUK has detailed datasets on dietary habits, which have been collected since inception of the registry. Data varies, and includes dietary indices on >5000 participants (e.g. Mediterranean Diet Score, Healthy Eating Index - 2010 and the Healthy Food diversity index) (Bowyer et al., 2018).

Dietary patterns, which are measured by category of foodstuff, have also been assessed through a food frequency questionnaire (FFQ) previously used in the EPIC Study (Bingham et al., 2008). For details of collection see Table 5.

Table 5. Food Frequency Questionnaire Collection

Date of Collection	N of individuals	Longitudinal data comments
1993-2001	4,470	2 or 3 time points in 890 individuals (Teucher et al., 2007)
2007	3,370	NA (one time point only)
2014-2018	5,440	2 time points (4 years apart) in 127 individuals

Note: NA= not applicable.

Socioeconomic data

The historical research focus of TwinsUK has shaped the main demographic of the twin cohort having middle socio-economic status and education typical of a volunteer group (Moayyeri et al., 2013; Steves et al., 2013). Socioeconomic status of the twin volunteers has been collected since the registry's inception through self-reported questions (e.g. highest educational qualification status). More recently the Index of Multiple Deprivation has been compiled for all volunteers having UK postal codes, and data are to be linked to national databases for retrospective and future collection.

- Index of multiple deprivation

Datasets from online government data repositories were combined, representing four of the United Kingdom's administrative countries: England (IMD version 2015) Scotland (IMD version 2016), Wales (IMD version 2014) and Northern Ireland (IMD version 2017). The IMD is a composite measure of area-level deprivation, and considers the following domains: income; employment; education, skills & training; health deprivation; crime; barriers to housing and services; and the living environment. As methods may vary between the countries, and ranks are inappropriate (given the differing numbers of administrative districts in each country), the decile score was combined as a relative measure of deprivation. Datasets were matched to postcodes or LSOA codes at 17,498 timepoints for 12,041 individuals. Mean IMD decile score (considering all time points) was 6.49.

Future Directions and Collaborations

Longitudinal and detailed clinical, biochemical, behavioural, socio-economic and deep omics (including multi tissue characterisation) of participants for nearly 30 years has provided a unique resource to study complex diseases and domains of healthy ageing in the TwinsUK population. These, in conjunction with novel dynamic testing at study visits and lifestyle intervention studies, offer a unique opportunity to explore personalised medicine. High quality data collection, database management, biological sample storage, and statistical quality control enhance the resource, In addition, a key strength of the resource lies in the highly engaged and loyal population, this evident from the high retention levels of participation across studies. Blood, urine, DNA and multiple tissue samples are available for future measurements; Online questionnaires and active engagement with our twin participants using text, email and social networking enables responsive and agile data collection. Our 'Volunteer Advisory Panel' are key on developing new strategy and governance of participants, informing on decisions about the ethics, practicalities and appropriateness of potential studies.

The TwinsUK registry has a history of numerous successful scientific collaborations, and we remain committed to providing the scientific community with access to the phenotype data from the 'TwinsUK Resource'. TwinsUK has an exemplary record for data sharing with over 800 data access requests, 150,000 samples shared to over 100 collaborators, and over 600 publications in the past six years. Detailed descriptions for researchers of data and samples are on the data access pages of the website (<http://www.twinsuk.ac.uk/data-access/cohortdata-description/>), here over 10,000 phenotypes can be searched. Longitudinal Population studies funding from the Wellcome Trust continues to fund the core functions of TwinsUK and opens up the resource to successful cross cohort collaborations. Over the next five years TwinsUK will integrate electronic health records into an enhanced deep tissue 'omics resource and continue dynamic phenotypic testing into clinical visits. In addition, we will extend the age range of the registry to include volunteer twins from birth to adulthood, thus opening up the resource to study unique twin gene environment interactions across the life-course.

New efficient broad consent will ensure that the communication with participating twins is ethical and proportionate. New annual sociological questionnaires will harmonise with ELSA (English Longitudinal Study of Ageing) & other LPS (1946/1958). We will also standardise mental health phenotypes between the complementary Twins Early Development Study (TEDS) such that, together, TwinsUK and TEDS cohorts will be an unparalleled twin resource across the life-course. These developments will ensure TwinsUK will be a unique global resource of longitudinal omics and twin research across the life-course, with immense potential for future scientific exploitation.

Acknowledgements

TwinsUK is core funded by the Wellcome Trust, and grants WT081878MA and WT202786/Z/16/Z contributed to the majority of data described. In addition, TwinsUK receives funding from Medical Research Council, European Union, the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.

References:

- Allum, F., Shao, X., Guénard, F., Simon, M. M., Busche, S., Caron, M., ... Zondervan, K. T. (2015). Characterization of functional methylomes by next-generation capture sequencing identifies novel disease-associated variants. *Nature Communications*.
<https://doi.org/10.1038/ncomms8211>
- Barrios, C., Zierer, J., Gudelj, I., Štambuk, J., Ugrina, I., Rodríguez, E., ... Menni, C. (2015). Glycosylation Profile of IgG in Moderate Kidney Dysfunction. *Journal of the American Society of Nephrology*. <https://doi.org/10.1681/asn.2015010109>
- Barrios, C., Zierer, J., Würtz, P., Haller, T., Metspalu, A., Gieger, C., ... Menni, C. (2018). Circulating metabolic biomarkers of renal function in diabetic and non-diabetic populations. *Scientific Reports*. <https://doi.org/10.1038/s41598-018-33507-7>
- Bell, C. G., Xia, Y., Yuan, W., Gao, F., Ward, K., Roos, L., ... Spector, T. D. (2016). Novel regional

- age-associated DNA methylation changes within human common disease-associated loci. *Genome Biology*. <https://doi.org/10.1186/s13059-016-1051-8>
- Bell, J. T., Loomis, A. K., Butcher, L. M., Gao, F., Zhang, B., Hyde, C. L., ... Spector, T. D. (2014). Differential methylation of the TRPA1 promoter in pain sensitivity. *Nature Communications*. <https://doi.org/10.1038/ncomms3978>
- Bell, Jordana T., Tsai, P. C., Yang, T. P., Pidsley, R., Nisbet, J., Glass, D., ... Deloukas, P. (2012). Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genetics*. <https://doi.org/10.1371/journal.pgen.1002629>
- Bingham, S. A., Welch, A. A., McTaggart, A., Mulligan, A. A., Runswick, S. A., Luben, R., ... Day, N. E. (2008). Nutritional methods in the European Prospective Investigation of Cancer in Norfolk. *Public Health Nutrition*. <https://doi.org/10.1079/phn2000102>
- Bowyer, R. C. E., Jackson, M. A., Pallister, T., Skinner, J., Spector, T. D., Welch, A. A., & Steves, C. J. (2018). Use of dietary indices to control for diet in human gut microbiota studies, 1–11. <https://doi.org/10.1186/s40168-018-0455-y>
- Buil, Alfonso; Brown, Andrew A.; Lappalainen, Tuuli; Vinuela, Ana; Davies, Matthew N.; Zheng Hou-Feng; Richards J, Brent; Glass Daniel; Small, Kerrin.S; Durbin, Richard, Spector, Timothy D.; Dermitzakis, E. T. (2015). Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. *Nature Genetics*. <https://doi.org/10.1038/ng.3162>
- Davies, M. N., Krause, L., Bell, J. T., Gao, F., Ward, K. J., Wu, H., ... Wang, J. (2014). Hypermethylation in the ZBTB20 gene is associated with major depressive disorder. *Genome Biology*. <https://doi.org/10.1186/gb-2014-15-4-r56>
- Grundberg, E., Meduri, E., Sandling, J. K., Hedman, Å. K., Keildson, S., Buil, A., ... Deloukas, P. (2013). Global analysis of dna methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *American Journal of Human Genetics*. <https://doi.org/10.1016/j.ajhg.2013.10.004>
- Grundberg, E., Small, K. S., Hedman, Å. K., Nica, A. C., Buil, A., Keildson, S., ... Spector, T. D. (2012). Mapping cis-and trans-regulatory effects across multiple tissues in twins. *Nature*

Genetics. <https://doi.org/10.1038/ng.2394>

Jackson, M A, Goodrich, J. K., Maxan, M. E., Freedberg, D. E., Abrams, J. A., Poole, A. C., ...

Steves, C. J. (2016). Proton pump inhibitors alter the composition of the gut microbiota. *Gut*, 65(5), 749–756. <https://doi.org/10.1136/gutjnl-2015-310861>

Jackson, Matthew A., Verdi, S., Maxan, M. E., Shin, C. M., Zierer, J., Bowyer, R. C. E., ... Steves, C.

J. (2018). Gut microbiota associations with common diseases and prescription medications in a population-based cohort. *Nature Communications*. <https://doi.org/10.1038/s41467-018-05184-7>

Jackson, Matthew A, Jackson, M., Jeffery, I. B., Beaumont, M., Bell, J. T., Clark, A. G., ... Steves, C.

J. (2016). Signatures of early frailty in the gut microbiota. *Genome Medicine*, 8(1), 8. <https://doi.org/10.1186/s13073-016-0262-7>

Jarrar, Z. A., Ward, K. J., Mangino, M., Cherkas, L. F., Gill, R., Gillham-Nasenya, I., ... & Spector, T.

D. (2018). Definitive Zygosity Scores in the Peas in the Pod Questionnaire is a Sensitive and Accurate Assessment of the Zygosity of Adult Twins. *Twin Research and Human Genetics*, 21(2), 146-154. <https://doi.org/10.1017/thg.2018.9>

Kurushima, Y., Tsai, P. C., Castillo-Fernandez, J., Couto Alves, A., El-Sayed Moustafa, J. S., Le Roy,

C., ... Bell, J. T. (2019). Epigenetic findings in periodontitis in UK twins: A cross-sectional study. *Clinical Epigenetics*. <https://doi.org/10.1186/s13148-019-0614-4>

Le Roy, C. I., Beaumont, M., Jackson, M. A., Steves, C. J., Spector, T. D., & Bell, J. T. (2017).

Heritable components of the human fecal microbiome are associated with visceral fat. *Gut Microbes*. <https://doi.org/10.1080/19490976.2017.1356556>

Long, T., Hicks, M., Yu, H. C., Biggs, W. H., Kirkness, E. F., Menni, C., ... Telenti, A. (2017).

Whole-genome sequencing identifies common-to-rare variants associated with human blood metabolites. *Nature Genetics*. <https://doi.org/10.1038/ng.3809>

Menni, C., Jackson, M. A., Pallister, T., Steves, C. J., Spector, T. D., & Valdes, A. M. (2017a). Gut

microbiome diversity and high-fibre intake are related to lower long-term weight gain. *International Journal of Obesity*. <https://doi.org/10.1038/ijo.2017.66>

Menni, Cristina, Gudelj, I., MacDonald-Dunlop, E., Mangino, M., Zierer, J., Bešić, E., ... Valdes, A.

M. (2018a). Glycosylation Profile of Immunoglobulin G Is Cross-Sectionally Associated with

- Cardiovascular Disease Risk Score and Subclinical Atherosclerosis in Two Independent Cohorts. *Circulation Research*. <https://doi.org/10.1161/CIRCRESAHA.117.312174>
- Menni, Cristina, Keser, T., Mangino, M., Bell, J. T., Erte, I., Akmačić, I., ... Valdes, A. M. (2013b). Glycosylation of immunoglobulin G: Role of genetic and epigenetic influences. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0082558>
- Menni, Cristina, Lin, C., Cecelja, M., Mangino, M., Matey-Hernandez, M. L., Keehn, L., ... Valdes, A. M. (2018b). Gut microbial diversity is associated with lower arterial stiffness in women. *European Heart Journal*. <https://doi.org/10.1093/eurheartj/ehy226>
- Menni, Cristina, Zhai, G., MacGregor, A., Prehn, C., Römisch-Margl, W., Suhre, K., ... Valdes, A. M. (2013a). Targeted metabolomics profiles are strongly correlated with nutritional patterns in women. *Metabolomics*. <https://doi.org/10.1007/s11306-012-0469-6>
- Menni, Cristina, Zierer, J., Pallister, T., Jackson, M. A., Long, T., Mohnney, R. P., ... Valdes, A. M. (2017b). Omega-3 fatty acids correlate with gut microbiome diversity and production of N-carbamylglutamate in middle aged and elderly women. *Scientific Reports*. <https://doi.org/10.1038/s41598-017-10382-2>
- Mills, M. C., Barban, N., & Tropf, F. C. (2018). The Sociogenomics of Polygenic Scores of Reproductive Behavior and Their Relationship to Other Fertility Traits. *RSF: The Russell Sage Foundation Journal of the Social Sciences*. <https://doi.org/10.7758/RSF.2018.4.4.07>
- Mills, M. C., & Rahal, C. (2018). A scientometric review of genome-wide association studies. *Communications Biology*. <https://doi.org/10.1038/s42003-018-0261-x>
- Moayyeri, A., Hammond, C. J., Valdes, A. M., & Spector, T. D. (2012). Cohort profile: TwinsUK and healthy ageing twin study. *International Journal of Epidemiology*, 42(1), 76-85. <https://doi.org/10.1093/ije/dyr207>
- Moayyeri, A., Hammond, C. J., Hart, D. J., & Spector, T. D. (2013). The UK adult twin registry (TwinsUK resource). *Twin Research and Human Genetics*, 16(1), 144–149. <https://doi.org/10.1017/thg.2012.89>
- Nag et al (2019) Genome-wide scan identifies novel genetic loci regulating salivary metabolite levels. In preparation

- Ni Lochlainn, M., Bowyer, R. C. E., & Steves, C. J. (2018). Dietary Protein and Muscle in Ageing People: The Potential Role of the Gut Microbiome and its Metabolome. *European Geriatric Medicine*. <https://doi.org/10.1007/s41999-018-0097-4>
- Roos, L., Sandling, J. K., Bell, C. G., Glass, D., Mangino, M., Spector, T. D., ... Bell, J. T. (2017). Higher Nevus Count Exhibits a Distinct DNA Methylation Signature in Healthy Human Skin: Implications for Melanoma. *Journal of Investigative Dermatology*. <https://doi.org/10.1016/j.jid.2016.11.029>
- Shin, S.-Y., Fauman, E. B., Petersen, A.-K., Krumsiek, J., Santos, R., Huang, J., ... Soranzo, N. (2014). An atlas of genetic influences on human blood metabolites. *Nature Genetics*. <https://doi.org/10.1038/ng.2982>
- Steves, C. J., Jackson, S. H. D., and Spector, T. D. (2013). Cognitive change in older women using a computerised battery: a longitudinal quantitative genetic twin study. *Behav. Genet.* 43, 468–479. doi: 10.1007/s10519-013-9612-z
- Teucher, B., Skinner, J., Skidmore, P. M. L., Cassidy, A., Fairweather-Tait, S. J., Hooper, L., ... MacGregor, A. J. (2007). Dietary patterns and heritability of food choice in a UK female twin cohort. *Twin Research and Human Genetics*. <https://doi.org/10.1375/twin.10.5.734>
- The UK10K Consortium, Walter, K., Min, J. L., Huang, J., Crooks, L., Memari, Y., ... Zhang, W. (2015). The UK10K project identifies rare variants in health and disease. *Nature*. <https://doi.org/10.1038/nature14962>
- Verdi, S., Jackson, M. A., Beaumont, M., Bowyer, R. C. E., Bell, J. T., Spector, T. D., & Steves, C. J. (2018). An Investigation Into Physical Frailty as a Link Between the Gut Microbiome and Cognitive Health. *Frontiers in Aging Neuroscience*. <https://doi.org/10.3389/fnagi.2018.00398>
- Verdi, S., ... Steves, C. J. (2019) Cohort Profile: TwinsUK. *Wellcome Open Res*. In preparation
- Wells, P. M., Williams, F. M., Matey-Hernandez, M. L., Menni, C., & Steves, C. J. (2019). RA and the microbiome: do host genetic factors provide the link? *Journal of Autoimmunity*. <https://doi.org/10.1016/j.jaut.2019.02.004>
- Yuan, W., Xia, Y., Bell, C. G., Yet, I., Ferreira, T., Ward, K. J., ... Spector, T. D. (2014). An integrated epigenomic analysis for type 2 diabetes susceptibility loci in monozygotic twins.

Nature Communications. <https://doi.org/10.1038/ncomms6719>

Zhang, W., Spector, T. D., Deloukas, P., Bell, J. T., & Engelhardt, B. E. (2015). Predicting genome-wide DNA methylation using methylation marks, genomic position, and DNA regulatory elements. *Genome Biology*. <https://doi.org/10.1186/s13059-015-0581-9>

Zierer, J., Jackson, M. A., Kastenmüller, G., Mangino, M., Long, T., Telenti, A., ... Menni, C. (2018). The fecal metabolome as a functional readout of the gut microbiome. *Nature Genetics*. <https://doi.org/10.1038/s41588-018-0135-7>